Auton Robot (2014) 37:157–167
DOI 10.1007/s10514-014-9383-2

Autonomous robotic capture of non-cooperative target using visual servoing and motion predictive control

Benoit P. Larouche · Zheng H. Zhu

Received: 15 June 2013 / Accepted: 3 January 2014 / Published online: 12 January 2014

Abstract This paper presents a framework for autonomous capture operation of a non-cooperative mobile target in a 3-dimensional workspace using a robotic manipulator with visual servoing. The visual servoing with an eye-in-hand configuration is based on motion predictive control using Kalman filter for the on-line state and parameter estimation of the target. A transitional decision making process is developed to guide the robotic manipulator between the different phases of the capture operation by employing a custom metric that translates visual misalignments between the end-effector and the target into a guidance measurement. These phases include the target acquisition and approach stage and the alignment and capture phase. Experiments have been carried out on a custom designed and built robotic manipulator with 6 degrees of freedom. The objective is to evaluate the performance of the proposed motion predictive control scheme for the autonomous capturing task and to demonstrate the robustness of the proposed control scheme in the presence of noise and unexpected disturbances in vision system, sensory-motor coordination and constraints for the execution in real environments. Experimental results of the visual servoing control scheme integrated with the motion predictive Kalman filter indicate the feasibility and applicability of the proposed control scheme. It shows that when the target motion is properly predicted, the tracking and capture performance has been improved significantly.

B. P. Larouche · Z. H. Zhu (✉)
Space Engineering Design Laboratory, Department of Earth and Space Science and Engineering, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada
e-mail: gzhu@yorku.ca

B. P. Larouche
e-mail: larouche.benoit@gmail.com

Keywords Autonomy · Robotic manipulator · Visual servoing · Trajectory tracking · Capture · Non-cooperative target · Kalman filter · Motion predictive control

1 Introduction

The capture of non-cooperative targets using autonomous robotic manipulators as shown in Roe et al. (2004) and Ogilvie et al. (2008) presents a set of unique and interesting challenges as described in Edan (1995), Fumagall et al. (2012), Slaughter et al. (2008), not the least of which is the autonomy problem described in great detail in Liu and Chopra (2011), Liang and Ma (2011). Mobile robotics brings together the fusion of sensing, decision making, and acting in Chung et al. (2011). In autonomous robotic tracking and capturing of a non-cooperative target, computer vision is exclusively used as the primary feedback sensor in these tasks to control the pose of end-effector with respect to the target, seen in Cretual and Chaumetter (2001), Fang et al. (2012) and Wong et al. (1996). Among all vision-based control schemes, the position-based visual servoing with an eye-in-hand camera configuration has the advantages of precise view of a target and the natural relationship of the end-effector's pose with respect to the target in the workspace. However, the pose estimation in this configuration is prone to camera calibration errors, target model inaccuracies, image noise, and unexpected disturbances. Therefore, the focus of this study will be on the impact of pose estimation in visual servoing where the relative pose between a camera and a target is used for real-time control of robot motion, seen with limited success in Fang et al. (2011), Zhang et al. (2012) and Ghadyok et al. (2012). Furthermore, autonomous grasping is a critical task, as shown in Wang and Xie (2009), Hsiao et al. (2011) and Ignakov et al. (2012), in robotic capture

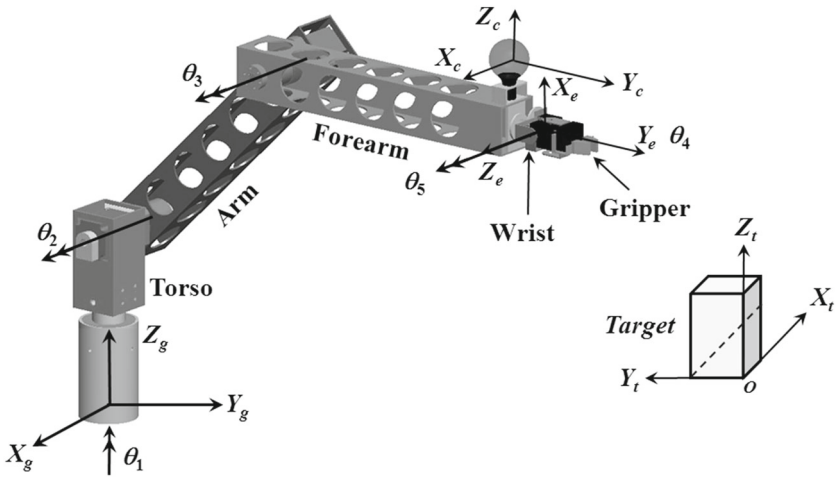

Fig. 1 Annotated robotic manipulator

operation and presents a change in the system kinematic and dynamic behavior without precise knowledge. To achieve the objective, an improved control scheme has been developed to obtain reliable pose and velocity estimates of a moving target from noisy image data using photogrammetry and Kalman filtering. The latter not only enhances the pose and velocity estimation from the noisy image data but more importantly provides smooth pose and rate estimates for the robotic control system when the vision system loses its tracking of the target momentarily. Once the target's pose and velocity are determined, the robotic trajectory to intercept the target at a rendezvous point will be planned and executed. The trajectory planning will be subjected to the constraints such as the maximum allowed velocity of the end-effector and the physical limits of the workspace and the joint ranges. A decision making metric has been developed and examined for a robotic manipulator attempting to approach and capture a non-cooperative target, where five critical alignment offsets are identified based on geometric character of the target and the existing work (two rotational offsets) as shown in Rekleitis et al. (2007). These offsets are then combined into a composite alignment index by a weighted summation for the total alignment measurement of the end-effector with respect to the the grasping point of a target. Individual weights are given depending on the stages of the capture operation in order to allow for a rapid approach, a high level of collision avoidance, and an increased rate of successful captures. The developed control scheme has been validated experimentally using a custom built robotic manipulator to lock on, track, approach and capture a moving target. The experimental results demonstrate a successful captures of a moving non-cooperative target with the proposed control scheme.

2 Kinematics of robotic manipulator

Consider an eye-in-hand (as defined in Fang et al. 2002), Pieper (see McKerrow 1991 for description) configured robotic manipulator as shown in Fig. 1. The actuators consist of three stepper motors attached to the first three links to control the torso (θ_1), shoulder (θ_2), and elbow (θ_3) joints, similar to the work of Schutter (1987). The orientation of the end-effector are controlled by two wrist joints actuated by two servo motors, namely, wrist rotation (θ_4) and wrist angle (θ_5) joints. The last joint is a translational joint that controls the gripper. Since the gripper will be activated after the end-effector is aligned with the grasping point of the target, we can consider only the degrees of freedom (DOF) of the first five joints that affect the pose of the end-effector in the robotic controller development, as shown in Kurniawati et al. (2012) and Klein and Huang (1983). The visual sensor in this work is a digital camera mounted next to the end-effector to monitor the target motion in a 3-dimensional (3D) workspace similar to Lippiello et al. (2008). The camera is assumed calibrated and the intrinsic and extrinsic parameters, such as the focal length, the physical size and resolution of image sensor, the transformation matrix between the camera and the end-effector, are known. However, the target can move in the 3D workspace and its pose and velocity are unknown in advance. Therefore, the robotic control problem can be defined as the following two tasks:

Task 1: Given a desired pose of the target in the workspace statically or dynamically, design a proper set of joint trajectories for the manipulator such that the end-effector can move to the desired grasp position and orientation smoothly with the minimum time.

Task 2: Once the end-effector is in the vicinity of the target, design a proper set of joint input for the manipulator to grasp the stationary or dynamic target without collision.

Let $\{X_g\}$ and $\{X_e\} \in \Re^3$ be the end-effector position vectors defined in the global and local coordinate systems in the 3D workspace and $\{\theta\} = \{\theta_1, \theta_2, \ldots, \theta_5\}^T \in \Re^5$ be the vector of corresponding joint angles defined in the joint space. The kinematics of a robotic manipulator defines the relationship between the joint positions and the corre-

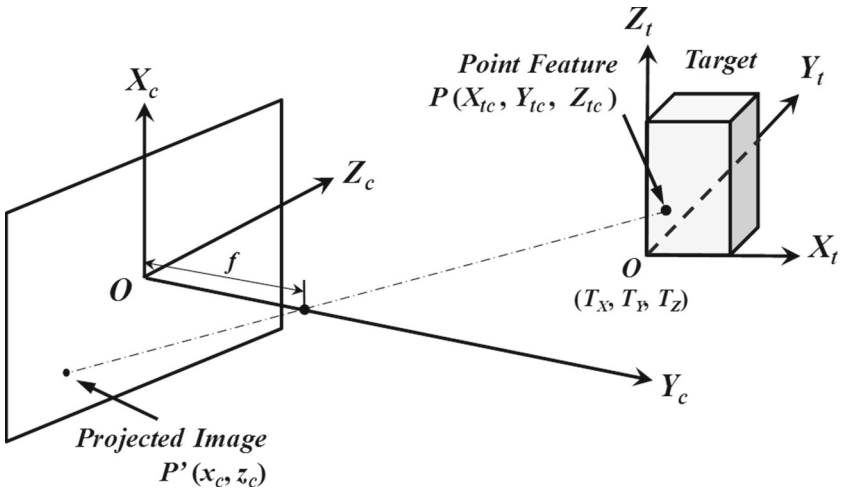

Fig. 2 Photogrammetric frames of reference

sponding position of the end-effector in the workspace, such that,

$$\{X_g\} = [T_{ge}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)]\{X_e\} \tag{1}$$

where $[T_{ge}]$ is the transformation matrix from local to global coordinate systems.

Similarly, the transformation from the camera frame $\{X_c\}$ to the global stationary frame $\{X_g\}$ can be expressed as,

$$\{X_g\} = [T_{gc}(\theta_1, \theta_2, \theta_3, X_c, Y_c, Z_c)]\{X_c\} \tag{2}$$

where (X_c, Y_c, Z_c) are the coordinates of origin of the camera frame in the frame $\{X_e\}$.

Then, the velocity of the end-effector can be found by taking the first order time derivation of (1),

$$\{\dot{X}_g\} = [J]\{\dot{\theta}\} \tag{3}$$

where $[J]$ is the Jacobian matrix of the robotic manipulator.

For the manipulator path planning, the inverse kinematics gives the joint angle velocity $\{\dot{\theta}\}$ in terms of the end-effector velocity,

$$\{\dot{\theta}\} = [\tilde{J}^{-1}]\{\dot{X}_g\} \tag{4}$$

where $[\tilde{J}^{-1}]$ is the pseudo-inverse of the Jacobian matrix in (3) if the matrix is non-square,

$$[\tilde{J}^{-1}] = ([J]^T[J])^{-1}[J]^T$$

3 Pose estimation

The pose information of the target in a 3D workspace is extracted from a 2D image space by photogrammetry (Fang et al. 2002; Mittal and Paragios 2004) based on the collinearity equations and a least squares solution. As shown in Fig. 2, the image coordinates of markers on the target can be established by collinearity equations, such as,

$$x_c = -f\frac{X_{tc}}{Y_{tc}}, \quad z_c = -f\frac{Z_{tc}}{Y_{tc}} \tag{5}$$

where (x_c, z_c) are the coordinates of a projected point on the camera's image plane containing only a 2D measurement, (X_{tc}, Y_{tc}, Z_{tc}) are the spatial coordinates of the marker with respect to the camera coordinate system, and f is the focal length of the camera. Here, it is implied that the focal length is much less than the distance between the marker and the camera.

The target point (X_{tc}, Y_{tc}, Z_{tc}) can be further decomposed into two parts:

$$\begin{Bmatrix} X_{tc} \\ Y_{tc} \\ Z_{tc} \end{Bmatrix} = [T_{tc}] \begin{Bmatrix} X_t \\ Y_t \\ Z_t \end{Bmatrix} + \begin{Bmatrix} X_{T0} \\ Y_{T0} \\ Z_{T0} \end{Bmatrix} \tag{6}$$

where $[T_{tc}]$ is the rotational matrix from the coordinate system attached to the target to the camera coordinate system, (X_t, Y_t, Z_t) is the coordinates of the marker in the local target coordinate system and (X_{T0}, Y_{T0}, Z_{T0}) is the coordinates of origin of the local target coordinate system in the camera coordinate system.

Substituting (6) into (5), we obtain:

$$\begin{Bmatrix} x_c \\ z_c \end{Bmatrix} = \begin{Bmatrix} -f\dfrac{r_{11}X_t + r_{12}Y_t + r_{13}Z_t + X_{T0}}{r_{21}X_t + r_{22}Y_t + r_{23}Z_t + Y_{T0}} \\ -f\dfrac{r_{31}X_t + r_{32}Y_t + r_{33}Z_t + Z_{T0}}{r_{21}X_t + r_{22}Y_t + r_{23}Z_t + Y_{T0}} \end{Bmatrix} \tag{7}$$

where r_{ij} are the functions of the Euler (roll–pitch–yaw) angles (Ω, Φ, Θ) of the target with respect to the camera's image plane. These equations are highly non-linear and involve six unknowns: the three rotational angles (Ω, Φ, Θ) and the three Cartesian coordinates (X_{T0}, Y_{T0}, Z_{T0}). They are the orientation and position of the target relative to the

Table 1 Test results of photogrammetry algorithm

Vision	Accuracy (%)	Error (%)	Precision (%)	Error (%)
X axis	93.12	±0.23	99.16	±0.05
Y axis	87.32	±0.75	97.01	±0.04
Z axis	94.50	±0.31	99.87	±0.05
Roll	99.45	±0.44	99.45	±0.08
Pitch	88.20	±0.80	88.45	±0.09
Yaw	88.13	±0.80	89.21	±0.08

camera coordinate system. In order to solve (7) mathematically, a minimum of three markers are required. In this work, four markers are used to (i) avoid the solution ambiguity and (ii) provide system redundancy so that the loss of one target point can be tolerated during the tracking process. The resulting equation is solved by the least-square method iteratively until the residual errors of the measurement satisfy a pre-set convergence criterion.

The photogrammetry algorithm is implemented together with the Open-sourced Computer Vision (OpenCV) library. The program outputs the coordinates of the markers on the image plane into the photogrammetry algorithm to calculate the 6 DOF of the target with respect to the camera. In order to improve the stability of the vision system, a Kalman filter embedded in the OpenCV is running in parallel with the photogrammetry algorithm to predict subsequent positions of the markers in concurrent images for the photogrammetry algorithm in case the vision system loses tracking of the markers, see Bradski and Kaehler (2008). Testing of the photogrammetry algorithm proceeded with a web cam of one mega-pixel resolution and the testing results are shown in Table 1. The accuracy of the algorithm is very high with the lowest value reading 87.32 %, a value which translates at the capture distance to less than 0.9 cm in the Y axis of camera frame. Once the pose of the target in the camera coordinate system is determined, it must be transformed into the global coordinate system (X_g) by (2) as input to the robot controller.

4 Kalman filter enhanced target tracking

In the eye-in-hand visual servoing, the pose estimated by the photogrammetry is noisy and prone to the residual vibrations of linkages of the manipulator due to the inherent mechanical flexibility and unexpected disturbances. To avoid the instability of robot control, the output of photogrammetry is treated by a Kalman filter to obtain the best stable and/or smooth pose estimation with the knowledge of the system that we have. In order to accomplish this, an unscented Kalman filter is employed with the state space shown in Table 2.

Let the $\{x\}$ be the state vector of the target by including the pose and its velocity and acceleration with respect to the camera coordinate system:

Table 2 Kalman filter state space

Symbol	Kalman filter symbol	Variable
X	X_1	X position
$\dot{X}$	X_2	X velocity
$\ddot{X}$	X_3	X acceleration
Y	X_4	Y position
$\dot{Y}$	X_5	Y velocity
$\ddot{Y}$	X_6	Y acceleration
Z	X_7	Y position
$\dot{Z}$	X_8	Y velocity
$\ddot{Z}$	X_9	Y acceleration
Θ	X_{10}	Roll angle
$\dot{\Theta}$	X_{11}	Roll velocity
$\ddot{\Theta}$	X_{12}	Roll acceleration
Φ	X_{13}	Pitch angle
$\dot{\Phi}$	X_{14}	Pitch velocity
$\ddot{\Phi}$	X_{15}	Pitch acceleration
Ω	X_{16}	Yaw angle
$\dot{\Omega}$	X_{17}	Yaw velocity
$\ddot{\Omega}$	X_{18}	Yaw acceleration

$$\{x\} = \{T_x, \dot{T}_x, \ddot{T}_x, T_y, \dot{T}_y, \ddot{T}_y, T_z, \dot{T}_z, \ddot{T}_z, \Theta, \dot{\Theta}, \ddot{\Theta}, \Phi, \dot{\Phi}, \ddot{\Phi}, \Omega, \dot{\Omega}, \ddot{\Omega}\} \tag{8}$$

Then, the Kalman filter model of the target in a discrete time form can be defined as follows

$$\{x\}_{k+1} = [A]\{x\}_k + [B]\{w\}_k \tag{9}$$

where the subscript $(k+1)$ denotes the states at time step $k+1$, $[A]$ is the state transition matrix, and $[B]$ is the disturbance transition matrix associated with the process noise vector $\{w\}_k$.

The 18×18 transition matrix $[A]$ is constructed by six 3×3 block diagonal sub-matrices, such as,

$$[a] = \begin{bmatrix} 1 & dt & dt^2/2 \\ 0 & 1 & dt \\ 0 & 0 & 1 \end{bmatrix}$$

while the disturbance transition matrix $[B]$ is an 18×6 sparse matrix with the following non-zero elements:

$$B_{3(i-1)+1,i} = \frac{dt^3}{6}, \quad B_{3(i-1)+2,i} = \frac{dt^2}{2}, \quad B_{3(i-1)+3,i} = dt$$

Here, dt is the sampling period and $i = 1, 2, \ldots, 6$. The process noise vector $\{w\}_k$ contains the jerks of the target and is assumed to obey a zero-mean white Gaussian distribution with its covariance, $[Q]_k$,

$$\{w\}_k = \{\dddot{T}_x, \dddot{T}_y, \dddot{T}_z, \dddot{\Theta}, \dddot{\Phi}, \dddot{\Omega}\}^T \sim N(0, [Q]_k) \tag{10}$$

Generally, the process noise covariance matrix $[Q]$ is difficult to determine in advance because of the non-cooperative nature of the target and unknown motion of the camera. In the current work, it is found that the following constant process noise covariance matrix works well after tuning the Kalman filter in experiments,

$$[Q] = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix} \times 10^{-6}$$

The measurement model of the Kalman filter is defined as

$$\{z\}_{k+1} = [H]\{x\}_k + \{v\}_k \tag{11}$$

where $\{z\}_{k+1} = \{T_z, T_y, T_z, \Theta, \Phi, \Omega\}^T$ is the measurement vector defined in Table 3. It is the output from the photogrammetry module. The matrix $[H]$ is a 6×18 sparse matrix with the following non-zero elements:

$$H_{1,1} = H_{2,4} = H_{3,7} = H_{4,10} = H_{5,13} = H_{6,16} = 1 \tag{12}$$

while the 6×1 vector $\{v\}_k$ is the measurement noise and obeys a zero-mean Gaussian distribution with its covariance $[R]_k$,

$$\{v\}_k \sim N(0, [R]_k) \tag{13}$$

The measurement noise covariance matrices $[R]_k$ is determined by tuning with real data of residual errors of the measurement from the photogrammetry algorithm in experiments using a non-linear regression analysis. For the given markers' pattern on the target and the vision configuration in the current study, it is found the following constant matrix is a good starting point for the Kalman filter,

$$[R] = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \times 10^{-6}$$

After the vision system starts, the $[R]$ matrix can be updated based on the output of the photogrammetry algorithm if needed.

The Kalman filter defined in this way is decoupled from the photogrammetry process to facilitate an easy implementation. Following a standard procedure, the Kalman filter recursively runs in two major steps as follow

(1) *Time update*

$$\{\hat{x}\}_{k+1,1} = [A]\{\hat{x}\}_{k,k} \tag{14}$$

$$[P]_{k+1,k} = [A][P]_{k,k}[A]^T + [Q]_k \tag{15}$$

Table 3 Kalman filter measurement space

Symbol	Kalman filter symbol	Variable
X	Y_1	X position
Y	Y_4	Y position
Z	Y_7	Y position
Θ	Y_{10}	Roll angle
Φ	Y_{13}	Pitch angle
Ω	Y_{16}	Yaw angle

(2) *Measurement update*

$$\{\hat{x}\}_{k+1,k+1} = \{\hat{x}\}_{k+1,k} + [K_g]_{k+1}(\{z\}_{k+1} - [H]\{\hat{x}\}_{k+1,k}) \tag{16}$$

$$[P]_{k+1,k+1} = [P]_{k+1,k} - [K_g]_{k+1}[H][P]_{k+1,k} \tag{17}$$

where the matrix $[K_g]$ is the Kalman gain at the time step $k+1$:

$$[K_g]_{k+1} = [P]_{k+1,k}[H]_{k+1}^T([H]_{k+1}[P]_{k+1,k}[H]_{k+1}^T + [R]_{k+1})^{-1}$$

5 Motion predictive control in autonomous robotic capture

The motion predictive control in autonomous robotic capture can be generally divided into two phases: (i) the tracking and approaching, and (ii) the capture.

5.1 Tracking and approaching

The target tracking and approaching involves identifying, locking on, and following a static or dynamic target in the 3D workspace by controlling the joint positions in the joint space. This includes the planning of the trajectory for the robot joint positions, the wrist position and the orientation of the end- effector to optimize capture potential. Assuming there is no obstacles between the end- effector and the target, a direct path is planned in the joint space, instead of workspace, in order to minimize the time taken for the end-effector to approach and align with the target. It should note that this may result in a curved trajectory of the wrist in the workspace.

Let us define the state errors in the position-based visual servoing by the relative position error of the wrist, $\{e\}$, and the orientation error, $\{\Phi\}$ of the gripper, such that,

$$\{e\} = \{X_g\} - \{X_g^d\} = \begin{Bmatrix} X_g \\ Y_g \\ Z_g \end{Bmatrix} - \begin{Bmatrix} X_g^d \\ Y_g^d \\ Z_g^d \end{Bmatrix} \tag{18}$$

$$\{\Phi\} = \{\alpha_e\} - \{\alpha_e^d\} = \begin{Bmatrix} \alpha_e \\ \alpha_e \\ \alpha_e \end{Bmatrix} - \begin{Bmatrix} \alpha_e^d \\ \alpha_e^d \\ \alpha_e^d \end{Bmatrix} \tag{19}$$

where $\{X_g\}$ and $\{X_g^d\}$ are the current and desired wrist positions in the global frame, $\{\alpha_g\}$ and $\{\alpha_g^d\}$ are the current and desired gripper orientations in the local coordinate system $\{X_e\}$.

Correspondingly, the state velocity state errors can be written as

$$\{\dot{e}\} = \{\dot{X}_g\} - \{\dot{X}_g^d\}; \{\dot{\Phi}\} = \{\dot{\alpha}_e\} - \{\dot{\alpha}_e^d\} \tag{20}$$

The desired wrist pose is obtained by the vision system via the photogrammetry algorithm and the Kalman filter.

For the dynamic tracking and approaching, the desired pose has to be modified so that the end-effector will approach the intercept position with proper orientation instead of the current position and orientation. This can be achieved by replacing the desired pose in (18) and (19) with the following,

$$\{\tilde{X}_g^d\} = \{X_g^d\} + \{\dot{X}_g^d\}(\Delta T + t_0) \tag{21}$$

$$\{\tilde{\alpha}_e^d\} = \{\alpha_e^d\} + \{\dot{\alpha}_e^d\}(\Delta T + t_0) \tag{22}$$

where t_0 is the time interval between two updates of the computer vision system and $\Delta T = \|\{e\}\|/\|\{\dot{X}_g^{max}\}\|$ is the nominal shortest remaining time from the current position to the target position, $\{\dot{X}_g^{max}\}$ is the maximum velocity limit imposed on the end-effector, and the velocities of the desired position $\{\dot{X}_g^d\}$ and orientation $\{\dot{\alpha}_e^d\}$ are estimated by the Kalman filter. As the end-effector approaches the target, the ΔT will become smaller and the estimation of the intercept pose will be more accurate.

Thus, the error in the joint space consists of two parts, resulting from the position and orientation errors,

$$\Delta\{\theta\} \approx [\tilde{J}^{-1}]\{e\} + [T_\phi]\{\Phi\}; \quad \Delta\{\dot{\theta}\} \approx [\tilde{J}^{-1}]\{\dot{e}\} + [T_\phi]\{\Phi\} \tag{23}$$

where $[T_\Phi]$ is the transformation matrix taking the angular misalignments from the local coordinate system $\{X_e\}$ to the joint space.

The path tracking of the end-effector will be achieved by a proportional and derivative controller due to its simplicity and reliability, such that,

$$\Delta\{\theta_c\} = K_p \Delta\{\theta\} + K_d \Delta\{\dot{\theta}\} \tag{24}$$

where (K_p, K_d) are the gains of proportional and derivative control, subject to the constraint of the maximum velocity of the end-effector,

$$\{\dot{\theta}\} \approx [\tilde{J}^{-1}]\{\dot{X}_g^d\} + [T_\Phi]\{\dot{\Phi}\} \leq [\tilde{J}^{-1}]\{\dot{X}_g^{max}\} \tag{25}$$

5.2 Capture

Once the end-effector is close to the target and prepares for a capture, a metric is defined to take into account misalignments and distances between the gripper and the grasping point in order to transition between phases as well as to decide when to effect a capture. These were modularized in order to scale their weightings throughout the individual phases of the operation. They are as follows:

$$M_1 = \sqrt{\Delta X_e^2 + \Delta Y_e^2 + \Delta Z_e^2} \tag{26}$$

$$M_2 = \sqrt{\Delta Y_e^2} \tag{27}$$

$$M_3 = \sqrt{(\alpha_x - \alpha_x^d)^2 + (\alpha_y - \alpha_y^d)^2 + (\alpha_z - \alpha_z^d)^2} \tag{28}$$

$$M_4 = Y_e tan(\alpha_x - \alpha_x^d) + Z_e \tag{29}$$

$$M_5 = Y_e tan(\alpha_z - \alpha_z^d) + X_e \tag{30}$$

where M_1 represents the misalignment in the non-critical total distance, M_2 the critical horizontal distance, M_3 the total orientation misalignment, M_4 and M_5 the horizontal and vertical orientations of the end-effector, respectively.

These five offsets are then combined into a composite alignment index for the total alignment measurement of the end-effector by a weighted sum,

$$h = \sum_{i=0}^{5} w_i M_i \tag{31}$$

where w_i are the weights that are tuned in the experiment.

The index is further conditioned by a sliding average to define a threshold logic function (TL) to control the capture action, such that,

$$TL = \begin{cases} j < 100 & \text{Initialization} \quad \text{Approach} \\ j \geq 100 & \begin{cases} 500 < h & \text{Approach} \\ 85 < h \leq 500 & \text{Approach and Alignment} \\ h \leq 85 & \text{Grasp} \end{cases} \end{cases} \tag{32}$$

where j is the number of vision measurements. Once the approach and alignment ($85 < h \leq 500$) of the end-effector is initiated, the weight for M_2 is increased to ensure a collision with the target does not occur.

6 Experimental validation

6.1 Experimental set-up

The experimental validation was carried out with the use of a custom designed, constructed, and programmed robotic manipulator as shown in Fig. 3. The experimental set up employed a visually low-noise target shown in Fig. 4 with

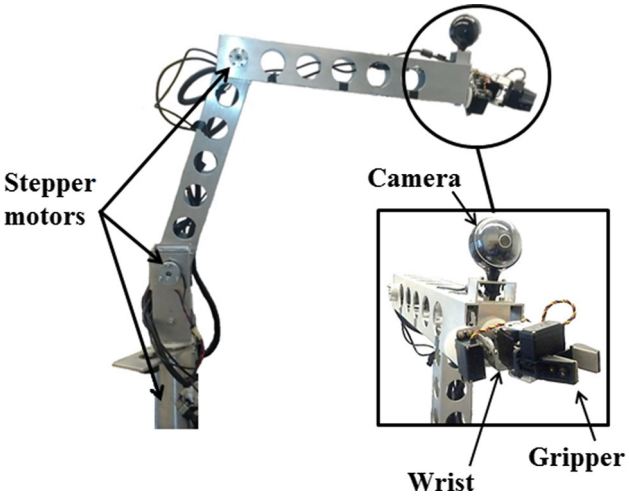

Fig. 3 Robotic manipulator

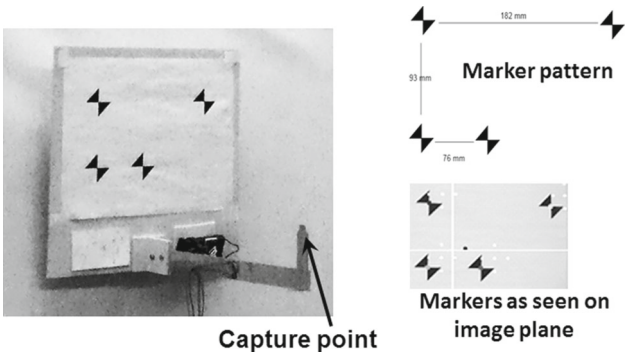

Fig. 4 Close up of target and pattern

four non-collinear markers in a pattern known and hard-coded into the photogrammetry algorithm. The pattern do not require any specific shape or position, only a high contrast.

The motion of the target is generated through the use of a single motor (referred as target motor in the following) that can run at different velocities. In addition, a pendulum anchor point can be moved, generating a modifiable arc motion for the target. An example of this motion is shown in Fig. 5 along with the robotic manipulator in the foreground. By varying the anchor points and keeping the knowledge of the path separate from the manipulator controller, we can generate several distinct capture scenarios for testing. Two testing scenarios were tested in this work, determined by the target velocity at capture as low and high velocity captures respectively. The upper limit for the low velocity capture was determined by increasing the velocity of the target motor until the point where, without motion prediction, the robotic manipulator was unable to track and capture the target. This velocity was established as 0.27 cm/s. The high velocity capture was performed at twice this velocity, namely, 0.54 cm/s.

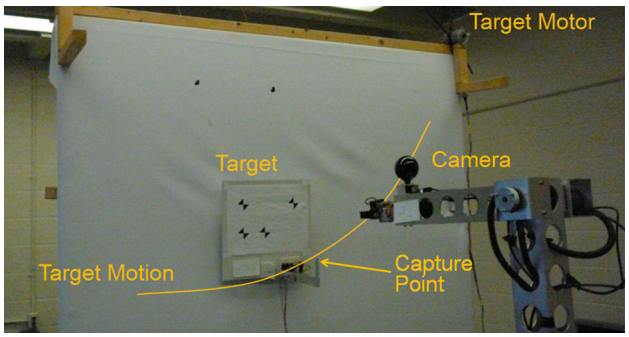

Fig. 5 Experimental set-up for validation

The capture follows several phases that are defined as follows.

1. Target acquisition
   (a) Markers finding
   (b) Target locking on and tracking
2. Approach
   (a) High speed approach to the vicinity of target
   (b) Collision avoidance approach before capture
3. Capture
   (a) Verify misalignments
   (b) Verify weighted summation
   (c) Activate capture

The transitions between the phases is a critical component of the operation and requires an autonomous solution capable of sensing the distance, the position, and the misalignment actively and deciding when the optimal transitions would occur. This will be illustrated in the following section.

6.2 Results and discussion

The experimental results of autonomous captures of a non-cooperative target will be presented in here to illustrate the desired and current joint positions of the torso (Θ_1), the shoulder (Θ_2), and the elbow (Θ_3). The capture operation will be examined in detail where the target is acquired in phase 1 and the predictive algorithm begins to output desired values for the joint angles. Phase 2 denotes the beginning of the tracking and the introduction of the actual joint positions. Phase 2a is the fast approach marked by the rapid changes in joint positions and large oscillations in the motion. Phase 2b is the near stage approach that is characterized by the slow changes in joint position and small oscillations in the motion due to an increased care in collision avoidance. Phase 3 is activated once the metric of misalignments denoted in Sect. 5.2 drops below the threshold value.

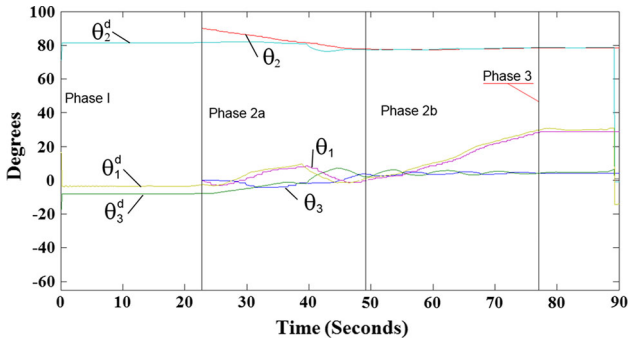

Fig. 6 Typical low velocity capture operation with the phases outlined and identified. Without Kalman filter

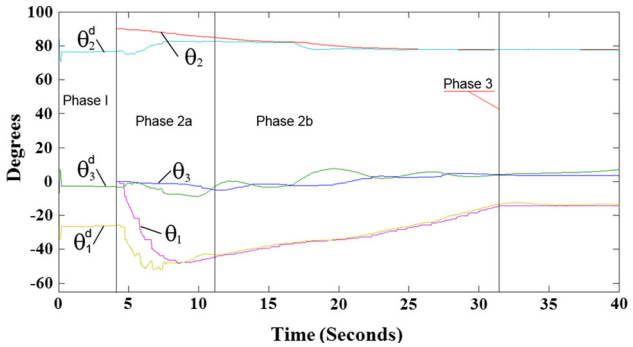

Fig. 7 Typical low velocity capture operation with the phases outlined and identified. With Kalman filter

6.2.1 Low velocity captures

Low velocity captures refer to the capture operation without predicting target motion by (21) and (22) in the controller. The target was moving at a linear velocity of $0.27\,cm/s$. Two cases were conducted to compare the impact of Kalman filter on the capture performance and the results are shown in Figs. 6 and 7. The differentiating factor between the two operations is that the Kalman filter was employed only in the second operation, in Fig. 7 and resulted in a much faster capture operation with smaller errors and fewer overshoots in joint positions. In terms of phase transitions, both operations perform similarly.

As shown in Figs. 6 and 7, Phase 1 was dominated with the vision system establishing a lock on the target and resolving the inverse kinematics. Obviously, the Kalman filter reduced the time required to perform this phase task significantly. Phase 2a emphasized the approaching speed, which resulted in a rapid motion in the end-effector of robotic manipulator. This can be best seen in the high velocity captures in the following section. Phase 2b started at the close proximity to the target and was executed at a slower end-effector velocity by emphasizing collision avoidance between the end-effector and the target. Of particular importance in this phase is to avoid any perturbations to the target by the end-effector. This

Table 4 Low velocity capture synopsis

Phase	Start time (s)	End time (s)	Elapsed time (s)
Without Kalman filter			
Phase 1	0	22.5	22.5
Phase 2a	22.5	48	25.5
Phase 2b	48	76.5	28.5
Phase 3	76.5	…	…
With Kalman filter			
Phase 1	0	4.2	4.2
Phase 2a	4.2	11.8	7.6
Phase 2b	11.8	31.9	20.1
Phase 3	31.9	…	…

is due to the fact that at the close proximity, the markers employed to locate the target in the 3D workspace take up a large percentage of the field of view of the camera and a quick acceleration of the target may result in the loss of lock to the markers and necessitate an abortion from the capture operation. Phase 3 is significant only in the sense that it marks the capture of the target and beginning of a new control regime aimed at decelerating the target, which will be a separate research subject. Table 4 provides a synopsis of the times where these events occur for the low velocity capture operations. Obviously, the Kalman filter enhanced significantly the capture performance in terms of shortening the operating time for a fast capture operation.

6.2.2 High velocity captures

The high velocity captures were executed when the target moved at a higher velocity of $0.54\,cm/s$. The target motion prediction in (21) and (22) must be considered in the controller in this case. In addition, the higher velocity leads to new challenges including a reduced time window to effect a capture operation due to the limited time the target spends within the operational field of robotic manipulator. Furthermore, high velocities incur higher errors in pose estimation and higher chances for collisions due to the system delay. Again, the comparisons between the cases with and without Kalman filter were performed.

Figure 8 shows the capture operation without using the Kalman filter and exhibits fairly similar trends to the low velocity captures from the previous section. The velocity slopes are steeper due to the faster joint speeds as would be expected.

To closely examine the behavior of the autonomous controller, we are going to take a look at a separate high velocity capture with the Kalman filter shown in Fig. 9 and its companion graph displaying the errors shown in Fig. 10. The errors of joint angles shown in Fig. 10 make it easy to distinguish

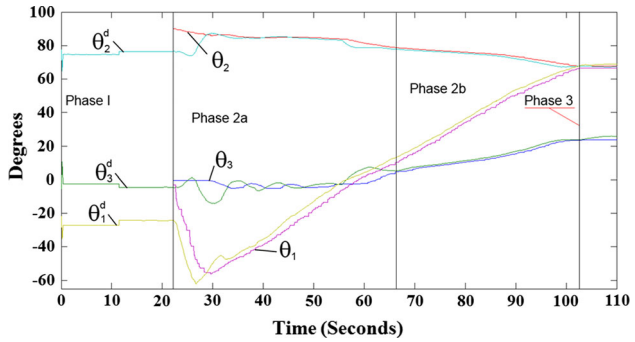

Fig. 8 Capture operation: Typical high velocity capture operation with the phases outlined and identified. Without Kalman filter

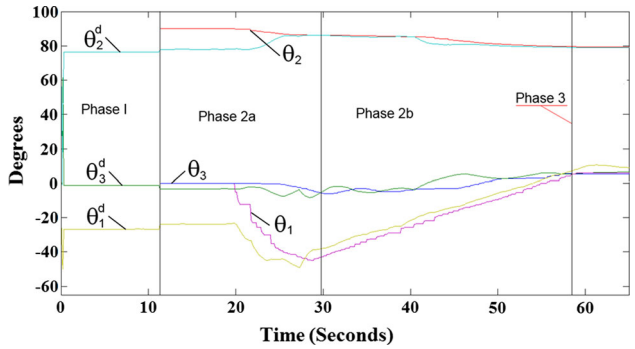

Fig. 9 Typical high velocity capture operation with the phases outlined and identified. With Kalman filter

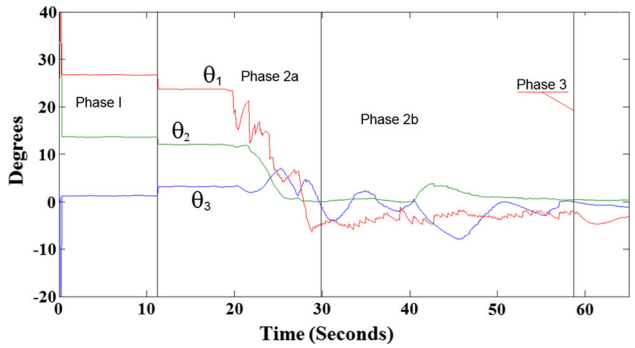

Fig. 10 Typical tracking errors in high velocity capture with the phases outlined and identified. With Kalman filter

the thresholds that trigger the different phases of the capture operation. The Phase 1 shows the vision system identifies and locks on the target only. No joint actuation is activated. The Phase 2a is characterized by the sharp reduction in joint angle errors, which corresponds to the fast approach of the end-effector to the target. Once in the close proximity to the target, the joint angle errors are reduced significantly and the Phase 2b starts until a successful capture is performed.

The lines in Fig. 10 represent the joint errors. In particular, the Θ_1 behaviour perfectly illustrates the separate behaviour profile between phase 2a and phase 2b. In phase 2a, the motion is jagged and indicates rapid changes in joint

Table 5 High vlocity cpture snopsis

Phase	Start time (s)	End time (s)	Elapsed time (s)
Without prediction			
Phase 1	0	22.5	22.5
Phase 2a	22.5	67.0	44.5
Phase 2b	67.0	102.3	35.3
Phase 3	102.3	…	…
With prediction			
Phase 1	0	11.0	11.0
Phase 2a	11.0	29.3	18.3
Phase 2b	29.3	58.1	28.8
Phase 3	58.1	…	…

positions that correlate to a rapid motion of the end-effector of robotic manipulator. Once the error value crosses into the near range operation mode, phase 2b, a much smoother curve emerges which signifies close proximity operations. This type of smooth phase transition provides an intuitive and bio-mimetic style capture that increases capture reliability while reducing collisions and jarring events. Table 5 provides a synopsis of the transitional times for the high velocity captures.

7 Conclusion

This paper investigated the autonomous capture of a non-cooperative target by a robotic manipulator using visual servoing and Kalman filter enhanced motion predictive control. The Kalman filter is used to enhance the real-time state and pose estimation of the target from the noisy image date and the trajectory planning for the dynamic capture of a non-cooperative target. The control of capture process has been divided into different phases and a transitional decision making process is developed to guide the robotic manipulator between the different phases of the capture operation using a custom metric that translates visual misalignments between the end-effector and the target into a guidance measurement. A custom 6 DOFs Pieper type robotic manipulator with eye-in-hand configuration is constructed to test and validate the performance of the proposed visual servoing, Kalman filter enhanced motion predictive control scheme for the autonomous capturing task and to demonstrate the robustness of the proposed control scheme in the presence of noise and unexpected disturbances in vision system and constraints for the manipulator operation in real environments. Experimental results demonstrated a successful phase transitioning strategy and capture operation. It also demonstrates the feasibility and applicability of the proposed control scheme in this paper.

Acknowledgments This work is supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Graduate Scholarships in Science and Technology (OGSST).

References

Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer Vision With the OpenCV Library*. Sebastopol: O'Reilly Media Inc.

Chung, T., Hollinger, G., & Isler, V. (2011). Search and pursuit-evasion in mobile robotics a survey. *Autonomous Robots*, *31*, 299–316.

Cretual, A., & Chaumetter, F. (2001). Application of motion-based visual servoing to target tracking. *The International Journal of Robotics Research*, *20*, 878–890.

Edan, Y. (1995). Design of an autonomous agricultural robot. *Applied Intelligence*, *5*, 45–50.

Fang, Y., Liu, X., & Zhang, X. (2011). Motion-estimation-based visual servoing of nonholonomic mobile robots. *IEEE Transactions on Robotics*, *6*, 1167–1175.

Fang, Y., Liu, X., & Zhang, X. (2012). Adaptive active visual servoing of nonholonomic mobile robots. *IEEE Transactions on Industrial Electronics*, *1*, 489–497.

Fang Y, Behal A, Dixon W, Dawson D (2002) "Adaptive 2.5d visual servoing of kinematically redundant robot manipulators". In: *Proceedings of the 41st IEEE Conference on Decision and Control, vol. 3*, (pp. 2860–2865), December 2002.

Fumagall, M., Ivaldi, S., Randazzo, M., et al. (2012). Force feedback exploiting tactile and proximal force/torque sensing. *Autonomous Robots*, *33*, 381–398.

Ghadyok, V., Goldin, J., & Ren, W. (2012). On the design and development of attitude stabilization, vision-based navigation, and aerial gripping for a low-cost quadrotor. *Autonomous Robots*, *33*, 41–68.

Hsiao, K., Kaelbling, L., & Lozano-Perez, T. (2011). Robust grasping under object pose uncertainty. *Autonomous Robots*, *31*, 253–268.

Ignakov, D., Okouneva, G., & Liu, G. (2012). Localization of a door handle of unknown geometry using a single camera for door-opening with a mobile manipulator. *Autonomous Robots*, *33*, 415–426.

Klein, C., & Huang, C. (1983). Review of pseudoinverse control for use with kinematically redundant manipulators. *IEEE Transactions on Systems*, *10*, 245–250.

Kurniawati, H., Bandyopadhyay, T., & Patrikalakis, N. (2012). Global motion planning under uncertain motion, sensing, and environment map. *Autonomous Robots*, *33*, 255–272.

Liang, J., & Ma, O. (2011). Angular velocity tracking for satellite rendezvous and docking. *Acta Astronautica*, *69*, 1019–1028.

Lippiello, V., Siciliano, B., & Villani, L. (2008). Interaction control of robot manipulators using force and vision. *International Journal of Optomechatronics*, *2*, 257–274.

Y.-C. Liu and N. Chopra, Semi-autonomous teleoperation in task space with redundant slave robot under communication delays. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 679–684), September 2011.

McKerrow, P. J. (1991). *Introduction to Robotics*. Sydney: Addison-Wesley.

A. Mittal and N. Paragios, Motion-based background subtraction using adaptive kernel density estimation. In: *Proceedings of the 2004 IEEE Computer on Computer Vision and Pattern Recognition, 2004, vol. 2*, (pp. II-302–II-309), June 2004.

A. Ogilvie, J. Allport, M. Hannah, et al., Autonomous robotic operations for on-orbit satellite servicing. In: *Proc. SPIE 6958, Sensors and Systems for Space Applications II*, no. 695809, April 2008.

Rekleitis, I., Martin, E., Rouleau, G., L'Archeveque, R., Parsa, K., & Dupuis, E. (2007). Autonomous capture of a tumbling satellite. *Journal of Field Robotics*, *24*(4), 275–296.

F. Roe, R. Howard, and L. Murphy, Automated rendezvous and capture system development and simulation for nasa. In: P. Motaghedi (Ed.), *Modeling, Simulation and Calibration of Space-Based Systems, Proceedings of the SPIE*.

De Schutter, J. (1987). A study of active compliant motion control methods for rigid manipulators based on a generic scheme. *IEEE International Conference on Robotics and Automation*, *4*, 1060–1065.

Slaughter, D., Giles, D., & Downey, D. (2008). Autonomous robotic weed control systems: A review. *Computers and Electronics in Agriculture*, *61*, 63–98.

Wang, H., & Xie, Y. (2009). Adaptive jacobian position/force tracking control of free-flying manipulators. *Robotics and Autonomous Systems*, *57*, 173–181.

Wong, A., Mayorga, R., Rong, L., & Liang, X. (1996). A vision based on-line motion planning of robot manipulators. In: Proceedings of the *IEEE/RSJ international conference on intelligent robots and systems - robotic intelligence interacting with dynamic worlds, vol. 1–3*, (pp. 940–948).

Zhang, S., Ding, Y., Hao, K., & Zhang, D. (2012). An efficient two-step solution for vision-based pose determination of a parallel manipulator. *Robotics and Computer-Integrated Manufacturing*, *2*, 182–189.

Benoit P. Larouche obtained his B.A.Sc. in Mechanical Engineering and M.A.Sc. in Aerospace Engineering from University of Toronto, and his Ph.D. from York University in Autonomous Robotics all located in Toronto, ON, Canada. His major field of study is satellites and autonomous robotics with a focus on on-orbit servicing and enabling technologies. He is currently a post-doctoral fellow in the Space Engineering Design Laboratory at York University. He concluded his master's degree with the launch of two nanosatellites, CanX-2 and NTS, back in 2008 which continue to be operational to this date. In addition, he's worked on several other currently operational nano-satellite missions that have since launched. He has published several papers on autonomous capture of non-cooperative targets as well as nano-satellite design. Dr. Larouche is a member of AIAA, IAC, and Prospectors & Developers Association. He is an Engineer in Training with the Professional Engineers Ontario.

Zheng H. Zhu received his B.Eng., M.Eng. and Ph.D. degrees in mechanics all from Shanghai Jiaotong University located in Shanghai, China. He also received his M.A.Sc. in robot control from University of Waterloo and Ph.D. in mechanical engineering from University of Toronto all located in Ontario, Canada. From 1993 to 1995, he worked as a research associate in Department of Mechanical and Industrial Engineering, University of Toronto. From 1995 to 2006, he was a senior engineer with the Curtiss–Wright Flow Control—Indal Technologies located in Mississauga, Ontario, Canada. Since 2006, he has been a professor with the Department of Earth and Space

Science and Engineering, York University in Toronto, Canada. He is the author of more than 100 articles. His research interests include on-orbit service robot, dynamics and control of tethered space system. He is the Editor-in-chief of the International Journal of Space Science and Engineering. Dr. Zhu is a senior member of IEEE and AIAA, member of ASME and CSME, and a licensed Professional engineer in Ontario.